# User-Defined Gestures for Holographic Medical Analytics

Tran Pham, Anthony Tang, Christian Jacob

University of Calgary

**ABSTRACT**

Advancements in augmented reality hardware help are helping to make the idea of "physically" interacting with 3D volumentric scans into a reality. In this poster, we explore how medical practitioners and technicians can explore and study 3D volumentric scans with head-mounted augmented reality technologies. We explore specifically how to design gestures for interacting with volumetric data given head-mounted AR. To inform this exploration, we designed and conducted a preliminary elicitation study, where eight participants created gestures to target resulting changes in the scene based on their expectation of how a working system would behave. Based on these gestures, we discuss how future gesture designers for head-mounted AR tools should explore interaction with 3D volumetric data.

**Keywords**: Augmented reality, hololens, gestural interaction, gestures, medical data

## 1 INTRODUCTION

The shift towards the development of touchless gestural interfaces there has been an increase of research on the applicability of these interfaces. One type of technology that use touchless interfaces are head mounted augmented reality displays. These head-tracked displays allow users to work in both real and virtual environments by displaying virtual content that is overlaid onto the real world. Users interact with the virtual content by means of hand gestures and basic controls.

This type of immersive environment presents interesting possibilities for medical imaging. Not only can these interfaces be used in sterile room environments, the holographic images allow for users to manipulate the data as they see fit. Finally, such immersive exploration may present new opportunities for teaching, and engagement for non-expert audiences.

The problem is that current interfaces use only a limited gesture for interaction, and that these gestures seem to be chosen primarily for ease of recognition (e.g. by cameras), rather than for ease of discoverability. For instance, the HoloLens currently only provides two different gestures: a "pinch" gesture, and a "push away" gesture. These gestures mimic, in turn, a mouse click, and a "back" action. Thus, the problem here is that we do not have a set of gestures for interacting with 3D data with regards to medical imaging—i.e. slicing, exploring and understanding volumetric data reconstructed from medical scans.

To guide our explorations in the design of discoverable and usable gestures for interacting with this volumetric data, we need grounds to base our explorations. To this end, the approach we take in this work is to use an elicitation study, where study participants are asked to create gestures (i.e. gestures are "elicited" from them). By eliciting gestures from non-technical users we can create a set of gestures composed of gestures people make without regard for recognition or technical concerns, i.e more intuitive. When referring to "intuitive" it is defined as coming naturally without excessive deliberation [2]. The possible expressive gestures that a person could use for a given interaction are vast, by eliciting gestures by means of a think-aloud protocol we can obtain insight on user mental models.



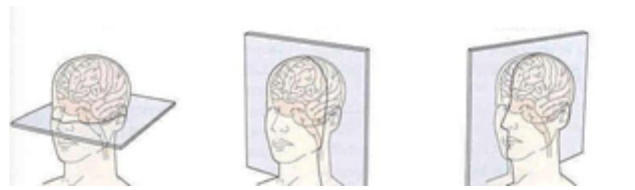

Figure 1: Several different possible cuts of a brain.

For example, cutting a 3D model of a head into two (Fig.1.) to examine different parts of the brain could be accomplished by many different gestures. For example, one could envision a chop motion across the hologram to split it, using a finger to slice the diagram, holding an object to cut the diagram (Fig.2.), jumping up two times and many more possibilities. Since the HoloLens works in an augmented reality environment it is possible that users may interact with the space with more than just their hands.

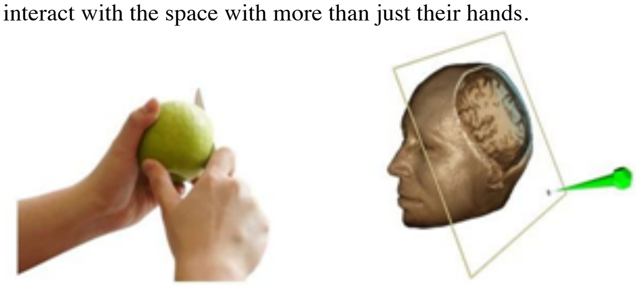

Figure 2: Possible mental models people could apply in arranging cut planes on a virtual model.

Based on our analysis of these gestures, we hope to contribute either a set of learnable/discoverable gestures, and principally, guidance to designers hoping to build gestures for AR.

## 2 RELATED WORK

Wobbrock et al. [3] proposed a method of gesture creation that involves having users as a part of the design process rather than having gestures defined by the system designers. In particular, they explore what surface gestures people make without regard for recognition or technical concerns. To investigate these idiosyncrasies they employed the guessability method that elicited gestures from non-technical users by presenting the effect from the gesture than asking users to create a gesture to match the interaction. They had users rate each gesture on a scale from 1-5 and then the gestures were analyzed. The result was a complete set of user-defined gestures for 27 different commands and taxonomy of surface gestures. Furthermore they found that one hand is preferred to two, that users rarely care about the number of fingers and that desktop idioms strongly influence user mental models.

This method elicitation of user-defined gestures is also employed by multiple related works and will be the main method we use to elicit the gestures. Wobbrock et al. also developed a four-dimensional taxonomy (form, nature, binding, and flow) for surface gestures in their work; this was then extended for touchless gesture by Piumsomboon et al.[1] by adding two more dimensions. The taxonomy used by Piusomboon is what we will be using to evaluate and classify the set of gestures. The taxonomy consists of 6 dimensions (form, nature, binding, flow, symmetry and locale) each dimension is comprised of different categories.

Bay et al. [4] employed the same guessability method as Wobbrock et al. [3], followed by two additional steps. After they elicited a complete gesture set from non-radiologists by means of paper prototyping they asked 2 professional radiologists to evaluate the applicability of the gesture set for professional use. They then implemented the gestures set into professional imaging solution to be evaluated by both professionals and non-professionals on three different sized displays.

A common finding between some of the elicitation studies is that most of the gestures elicited were physical and that one hand was preferred to two hands. Also by comparing the two gesture sets there were many similarities in the elicited gestures for certain tasks, indicating that there is a consistency with user-defined gestures. Much like the related work I look to contribute a set of user-defined gestures as a beginning point for future work and for consistency in the development of gestures for head-tracked displays.

Piumsomboon et al. create a set of user-defined gestures for augmented reality. They aimed to contribute to more consistent user-centered designed gestures in AR. Here, the authors argue that better control can be achieved by manipulations of the dynamical constraints rather than physical-based interaction. They suggest that hands should be treated as translucent rather opaque due to the fact that it could hinder the users experience when virtual objects are smaller than the user's hand. They also mention that there should be visual feedback to inform users of the contacts points between hands and objects. Another point they mention is that size does matter when it comes to the object size. Finally, they argued that is it easier to gesture an action when the user communicates as how they habitually performed the action or when they communicate it as an instructions. ("This is how I do it" vs. "this is how you should do it"). They found that gesturing how users would habitually perform an action was easier and more intuitive than gesturing an instruction.

## 3 Pilot Study

Informed by this prior work, we designed a study where participants were exposed to 16 referents based on micro-tasks that technicians would employ in exploring 3D volumetric medical data. These referents included, for example, zooming in and out of a data set, slicing into a data set in various ways, and manipulating (rotating, moving) the volume. In each case, participants were exposed to a "before" and "after" of each referrant, viewing both through the HoloLens. Participants were then asked to rate the quality and interpretability of the created gestures.

We recruited a total of ten participants to our pilot study. Nine were males and the average age was 22.8 years ($sd$ = 4.80). Only one participant had advanced prior experience with the HoloLens the others all had no prior experience. Of these nine are students and one a professor. Participants field of study included computer science, sociology, chemistry, psychology, and business.

### 3.1 Referents

We used a set of 16 referents: Uniform Scale, Scale x-axis, Scale y-axis, Scale z-axis, Roll (rotate in x-axis), Pitch (rotate in y-axis), Yaw (rotate in z-axis), Place plane, Turn plane x-axis, Turn plane y-axis, Move ball, Move plane, Cut ball, Move plane diagonal, Add a second plane. These corresponded to our understanding of how medical practitioners make use of volumetric data. For instance:

[Place Plane] – Participants started with a sphere and were asked to place a vertical "cut plane" in the middle of the sphere.

[Move ball] – Participants started with a sphere with a vertical cut plane and were asked to move just the ball.

[Move plane] – Participants started with a sphere with a vertical cut plane and were asked to move just the plane.

[Cut ball] – Participants started with a sphere with a vertical cut plane and were asked to cut either the right of left side of the plane.

[Move plane diagonal] – Participants started with a sphere with a vertical cut plane and were asked to move the plane diagonally.

[Add a second plane] – Participants started with the diagonal plane and were asked to add a second plane vertical cut plane.

Figure 3 illustrates one of these.

## 4 Discussion

Analysis of the gestures is still ongoing. At a preliminary level, however, we have found that for the referents 1-7 there was a lot of overlap for the gestures the participants chose but for referents 8-15 there was rarely any overlap. Often user mental models for referents 1-7 were influenced by the surface gestures they would use for the same interaction on their cellphones. When it came to the tasks involving the sphere and the plane (8-15) participants often asked more questions and took longer to produce a gesture since these tasks were unlike tasks that they are familiar with. Furthermore, with these tasks there was rarely any overlap with the gestures that participants chose.

Participants in the field of computer science were often concerned about the ability of the system to recognize the gestures and mentioned that they would imagine that there would be a menu on screen of some sort with to change between tools to execute the different action (add plane, move plane, cut sphere and rotate plane/sphere). But as for the participants not in the field of computer science they did mention concern for implementation and did not really imagine there being a menu but more of a legend on screen showing them what gestures they would need to do to perform certain actions.

To add to the discrepancy between the mental models of computer scientist and non-computer scientist, some of the computer scientist thought of having an extra physical tool to aid in the execution of the tasks involving the sphere and plane (8-15). And the majority imagined there being a menu to select from. Whereas non-computer scientist imagined being able to execute every tasks using only hand gestures and not needing anything on screen or physical.

## References


[1] T. Piumsomboon, A. Clark, M. Billinghurst and A. Cockburn. "User-defined gestures for augmented reality" in Proc. Extended Abstractson Human Factors in Computing Systems, 2013, pp. 955-960.

[2] S.A. Grandhi, G. Joue and I. Mittleberg. "Understanding naturalness and intuitiveness in gesture production: insights for touchless gestural interfaces" in Proc. SIGCHI Conference on Human Factors in Computing Systems, 2011, pp. 821-824.

[3] J.O. Wobbrock, M.R. Morris and A.D. Wilson. "User-defined gestures for surface computing," in Proc. SIGCHI Conference on Human Factors in Computing System, 200, pp.1083-1092.

[4] S. Bay, P. Brauner, T.Gossler and M. Ziefle, "Intuitive gestures on multi-touch displays for reading radiological images", in Proc. International Conference on Human Interface and the Management of Information, 2013, pp.22-31.